

Kapitel 3: Der *t*-Test

Durchführung eines <i>t</i> -Tests für unabhängige Stichproben	1
Durchführung eines <i>t</i> -Tests für abhängige Stichproben	5
Durchführung eines <i>t</i> -Tests für eine Stichprobe	8
Vertiefung: Vergleich <i>t</i> -Test für unabhängige und abhängige Stichproben in R	9

Durchführung eines *t*-Tests für unabhängige Stichproben

Dieser Abschnitt zeigt die Durchführung des in Kapitel 3.1 vorgestellten *t*-Tests für unabhängige Stichproben mit R. Das Beispiel beschäftigte sich mit der Frage, wie der Unterschied zwischen den Verarbeitungsbedingungen „strukturell“ und „bildhaft“ in der Erinnerungsleistung zu Stande gekommen ist. Handelt es sich vermutlich um einen zufälligen Unterschied, oder gibt es Grund zu der Annahme, dass sich die beiden Gruppen systematisch unterscheiden?

Das Experiment beinhaltete insgesamt drei unterschiedliche Verarbeitungsbedingungen. Mit den bislang vorgenommenen Angaben weiß R noch nicht, zwischen welchen Gruppen es den Unterschied auf Signifikanz testen soll. Daher müssen die beiden betreffenden Gruppen definiert werden. Die Verarbeitungsbedingung ist für jede Versuchsperson eindeutig festgelegt.

Mithilfe der Funktion `subset()` können Sie den relevanten Teil des Datensatzes auswählen. Mit dem Symbol „|“ wird angegeben, dass nur die Fälle ausgewählt werden sollen, die entweder in der Bedingung „strukturell“ oder in der Bedingung „bildhaft“ sind.

```
strk.bild <- subset(beispiel,
                  beispiel$bed == "strukturell"
                  | beispiel$bed == "bildhaft")
```

Verschaffen Sie sich als nächstes einen Überblick über ihre Daten:

```
library(psych)
describeBy(strk.bild$ges, strk.bild$bed)
Descriptive statistics by group
group: strukturell
  vars  n mean  sd median trimmed  mad min max range skew kurtosis  se
x1    1 50  7.2 3.16    7   7.03 2.97  1 15  14 0.45  -0.08 0.45
-----
group: bildhaft
  vars  n mean  sd median trimmed  mad min max range skew kurtosis  se
x1    1 50  11 4.14   10  10.75 4.45  5 26  21 0.9   1.51 0.59
```

<https://lehrbuch-psychologie.springer.com/content/zusatztexte-mit-anleitungen-zu-spss-r-und-gpower-sowie-datensätze>

Aus: Rasch, Friese, Hofmann & Naumann (2021). *Quantitative Methoden. Band 1*, 5. Auflage. Heidelberg: Springer.

© Springer-Verlag GmbH Deutschland, ein Teil von Springer Nature

```
-----  
-----  
group: emotional  
NULL
```

Hier sehen Sie unter anderem die Anzahl der Versuchspersonen in jeder der beiden beteiligten Gruppen, die jeweiligen Mittelwerte, Standardabweichungen und Standardfehler des Mittelwertes. Alle diese Konzepte sollten Ihnen aus den Kapiteln 1-3 vertraut sein.

Als nächstes führen wir den *t*-Test mit der Funktion `t.test()` durch. Zunächst wird die abhängige Variable, also die Gesamtzahl erinnelter Adjektive, angegeben und danach die unabhängige Variable.

```
t.unab <- t.test(strk.bild$ges ~ strk.bild$bed)
```

Geben Sie `t.unab` in der Konsole ein und Sie erhalten folgendes Ergebnis:

```
      Welch Two Sample t-test  
  
data:  strk.bild$ges by strk.bild$bed  
t = -5.1575, df = 91.653, p-value = 0.000001437  
alternative hypothesis: true difference in means is not equal to 0  
95 percent confidence interval:  
 -5.263399 -2.336601  
sample estimates:  
mean in group strukturell      mean in group bildhaft  
                7.2                11.0
```

Hinweis: Falls Sie den *p*-Wert nicht in der wissenschaftlichen Notation angezeigt haben möchten, geben Sie `options(scipen=999)` ein und lassen Sie sich das Ergebnis erneut ausgeben.

R führt per Standardeinstellung einen Welch-Test durch, der keine gleichen Varianzen zwischen den Gruppen annimmt und die Freiheitsgrade entsprechend korrigiert. Für den Fall, dass Varianzhomogenität vorliegt, müssen Sie das Argument `var.equal = TRUE` ergänzen.

```
t.unab.var <- t.test(strk.bild$ges ~ strk.bild$bed,  
                    var.equal = TRUE)
```

Geben Sie `t.unab.var` in der Konsole ein und Sie erhalten folgendes Ergebnis:

```
      Two Sample t-test  
  
data:  strk.bild$ges by strk.bild$bed  
t = -5.1575, df = 98, p-value = 0.000001305  
alternative hypothesis: true difference in means is not equal to 0  
95 percent confidence interval:  
 -5.262133 -2.337867
```

<https://lehrbuch-psychologie.springer.com/content/zusatztexte-mit-anleitungen-zu-spss-r-und-gpower-sowie-datensätze>

Aus: Rasch, Friese, Hofmann & Naumann (2021). *Quantitative Methoden. Band 1*, 5. Auflage. Heidelberg: Springer.

```
sample estimates:
mean in group strukturell      mean in group bildhaft
                7.2                11.0
```

Führen Sie einen Levene-Test durch, um zu überprüfen, ob die Varianzen in beiden Gruppen gleich sind. Dies ist eine Voraussetzung für die Durchführung des t -Tests. Aus den deskriptiven Statistiken können wir entnehmen, dass sich die Standardabweichungen deskriptiv unterscheiden (Die Varianz ist das Quadrat der Standardabweichung, siehe Kapitel 1.3). Ist dieser Unterschied so groß, dass die Varianzen als nicht homogen gelten müssen? Der Levene-Test nutzt die F -Verteilung (siehe Kapitel 5), um diese Frage zu überprüfen.

Für die Durchführung des Levene-Tests benötigen Sie das Paket `car`. Installieren und laden Sie es.

```
install.packages("car")
library(car)
```

Der Levene-Test wird mit dem Befehl `leveneTest()` durchgeführt. Als erstes wird die abhängige Variable und als zweites die unabhängige Variable angegeben.

```
leveneTest(strk.bild$ges, strk.bild$bed, center = mean)
```

```
Levene's Test for Homogeneity of Variance (center = mean)
```

```
      Df F value Pr(>F)
group 1  3.7638 0.05525 .
      98
```

```
---
```

```
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Der p -Wert zeigt ein marginal signifikantes Ergebnis an. Wir müssen also davon ausgehen, dass die Varianzen tendenziell inhomogen sind. Das bedeutet, dass wir bei der Frage der Signifikanz des Ergebnisses besonders auf die Werte von `t.unab` achten werden. Einige Autoren empfehlen, auch unabhängig vom Ergebnis des Levene-Tests grundsätzlich das t -Test Ergebnis korrigierter Varianzen zu verwenden, also die Zeile „Varianzen sind nicht gleich“ (Delacre, Lakens, & Leys, 2017). Dieser Empfehlung schließen wir uns an.

Die Resultate für den t -Test können Sie sich also in zwei Versionen ausgegeben lassen: für den Fall der Annahme gleicher Varianzen und für den konträren Fall. In beiden Fällen finden wir natürlich identische Werte für die mittlere Differenz von $-3,8$. Aus den leicht unterschiedlichen Konfidenzintervallen können wir schließen, dass sich die Standardfehler unterscheiden (zur Berechnung von Konfidenzintervallen siehe Kapitel 2.3).

Auch der empirische t -Wert von $5,16$ ist in beiden Fällen angenommener gleicher und ungleicher Varianzen identisch. Dieser Wert entspricht dem in Kapitel 3.1 „per Hand“ ermittelten t -Wert. Pro Gruppe gibt es einen Freiheitsgrad weniger als Versuchspersonen. Im Fall der Annahme ungleicher Varianzen korrigiert R diese Freiheitsgrade nach unten. Auf diese Weise liefert ein t -Test noch immer zuverlässige Ergebnisse, selbst wenn die Gleichheit der Varianzen nicht gewährleistet sein

<https://lehrbuch-psychologie.springer.com/content/zusatztexte-mit-anleitungen-zu-spss-r-und-gpower-sowie-datensätze>

Aus: Rasch, Friese, Hofmann & Naumann (2021). *Quantitative Methoden. Band 1*, 5. Auflage. Heidelberg: Springer.

© Springer-Verlag GmbH Deutschland, ein Teil von Springer Nature

sollte. Eine Freiheitsgradkorrektur kann die Auftretenswahrscheinlichkeit eines empirischen t -Werts unter Annahme der Nullhypothese entscheidend beeinflussen. Im vorliegenden Fall macht dies allerdings keinen Unterschied. In beiden Fällen ist das Ergebnis hoch signifikant. Die Interpretation lautet also: Es ist sehr unwahrscheinlich, dass die beiden Mittelwerte der Gruppen „strukturell“ und „bildhaft“ Populationen mit demselben Mittelwert entspringen. Die Nullhypothese wird verworfen und die Alternativhypothese angenommen. Offenbar gibt es einen systematischen Unterschied zwischen den beiden Gruppen, der die Mittelwertdifferenz mit verursacht hat.

Stellen Sie sich vor, die Wahrscheinlichkeit für den empirischen t -Wert wäre nicht so klein gewesen, sondern läge bei $p = 0,08$. In diesem Fall wäre das Ergebnis nach dem konventionellen Kriterium von $\alpha = 0,05$ nicht signifikant bei der von R durchgeführten zweiseitigen Testung. Da wir aber die Richtung der Differenz zwischen den Gruppen durch eine Theorie vorhergesagt haben (siehe Einleitung zu Band 1), ist eine einseitige Testung legitim. In diesem Fall dürften wir den p -Wert halbieren. Mit $p = 0,04$ wäre dieser bei einseitiger Testung signifikant (siehe Kapitel 3.2.3).

<https://lehrbuch-psychologie.springer.com/content/zusatztexte-mit-anleitungen-zu-spss-r-und-gpower-sowie-datensätze>

Aus: Rasch, Friese, Hofmann & Naumann (2021). *Quantitative Methoden. Band 1*, 5. Auflage. Heidelberg: Springer.

© Springer-Verlag GmbH Deutschland, ein Teil von Springer Nature

Durchführung eines *t*-Tests für abhängige Stichproben

Achtung: Für diesen Abschnitt ist die Datei „Messwiederholung.sav“ relevant, die Sie auf der Internetseite zum Download finden.

Lesen Sie zunächst den Datensatz ein:

```
library(foreign)
messw <- read.spss("Messwiederholung.sav", to.data.frame = TRUE)
```

Der *t*-Test für abhängige Stichproben ist das geeignete statistische Verfahren, wenn Daten aus Untersuchungen mit Messwiederholung analysiert werden sollen. Messwiederholung bedeutet in diesem Fall, dass eine Person zwei Mal Werte auf derselben abhängigen Variable geliefert hat.

In Kapitel 3.5 haben Sie das Beispiel einer motorischen Aufgabe gelesen, die alle Versuchspersonen zwei Mal durchgeführt haben. In der Aufgabe mussten die Versuchspersonen innerhalb von 30 Sekunden möglichst häufig eine kurze Sequenz mit den Fingern tippen. Die Anzahl richtiger Sequenzen innerhalb dieser Zeit bildete die abhängige Variable. Mit dem *t*-Test für abhängige Stichproben lässt sich überprüfen, ob die Versuchspersonen mit der zweiten Durchführung dieser Aufgabe eher besser geworden sind (z.B. durch einen Übungseffekt) oder eher schlechter (z.B. durch Motivationsverlust). Wir wollen nun dieses Beispiel mit R nachvollziehen. In unserem Beispiel handelt es sich um die Messungen 1 und 2. Verschaffen Sie sich zunächst einen Überblick und lassen Sie sich die deskriptiven Statistiken und die Korrelation ausgeben:

```
describe(data.frame(messw$Messung1, messw$Messung2))
```

	vars	n	mean	sd	median	trimmed	mad	min	max	range	skew	kurtosis	se
messw.Messung1	1	36	16.56	4.93	17.0	16.53	4.45	4	28	24	-0.05	0.36	0.82
messw.Messung2	2	36	17.28	5.22	16.5	17.20	3.71	5	29	24	0.14	0.38	0.87

Die Korrelation berechnet sich mit der Funktion `cor()`. Geben Sie die beiden Variablen ein. Das Argument `use` spezifiziert, wie mit fehlenden Werten in den Daten umgegangen werden soll. Das Attribut `pairwise.complete.obs` gibt an, dass Fälle mit fehlenden Werten bei den beiden Variablen ausgeschlossen werden sollen.

```
cor(messw$Messung1, messw$Messung2, use = "pairwise.complete.obs")
[1] 0.6611618
```

Der *t*-Test für abhängige Stichproben wird ebenfalls mit der Funktion `t.test()` durchgeführt. Geben Sie die beiden Variablen an und danach müssen Sie das Argument `paired = TRUE` ergänzen, damit R die Daten als abhängige behandelt.

```
t.abh <- t.test(messw$Messung1, messw$Messung2, paired = TRUE)
```

Es resultiert dieser Output, wenn Sie `t.abh` in der Konsole eingeben:

<https://lehrbuch-psychologie.springer.com/content/zusatztexte-mit-anleitungen-zu-spss-r-und-gpower-sowie-datensätze>

Aus: Rasch, Friese, Hofmann & Naumann (2021). *Quantitative Methoden. Band 1*, 5. Auflage. Heidelberg: Springer.

© Springer-Verlag GmbH Deutschland, ein Teil von Springer Nature

Paired t-test

```
data: messw$Messung1 and messw$Messung2
t = -1.0353, df = 35, p-value = 0.3077
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 -2.1384816  0.6940372
sample estimates:
mean of the differences
      -0.7222222
```

In den deskriptiven Statistiken können Sie erkennen, wie viele Sequenzen die Versuchspersonen zum ersten bzw. zweiten Messzeitpunkt durchschnittlich geschafft haben. Daraus können Sie bereits eine erste subjektive Einschätzung vornehmen, ob und in welcher Richtung sich die beiden Messzeitpunkte unterscheiden. Dies ist sehr wichtig. Der Signifikanzwert am Ende des *t*-Tests sagt lediglich etwas über die statistische Bedeutsamkeit einer Differenz zwischen zwei Variablen aus. Über die Richtung dieser Differenz enthält er keine Information.

Als nächstes wurde die bivariate Korrelation zwischen den beiden Messwertreihen berechnet. Die Korrelation ist ein Maß für den Zusammenhang zweier Variablen oder für die Frage, wie ähnlich sich diese Variablen sind. In diesem Fall lässt sich das also in die Frage übersetzen, wie ähnlich die Performanz der Versuchspersonen zu Messzeitpunkt 1 und Messzeitpunkt 2 war. In Kapitel 4 erfahren Sie mehr über das Konzept der Korrelation.

In dem Output des *t*-Tests erkennen Sie zunächst den Wert von $-0,722$ für die mittlere Differenz zwischen den getesteten Variablen. Dies entspricht dem Betrag, den Sie erhalten, wenn Sie in den deskriptiven Statistiken die beiden Mittelwerte voneinander subtrahieren (abgesehen von Rundungen). Zum zweiten Messzeitpunkt haben die Versuchspersonen also etwas häufiger die Sequenz tippen können als zum ersten Messzeitpunkt.

Schließlich sehen Sie das 95% Konfidenzintervall um die mittlere Differenz. Am Konfidenzintervall können Sie bereits erkennen, dass es um Null variiert und positive wie negative Werte einschließt. Dies ist ein weiterer Hinweis darauf, dass die gefundene Differenz vermutlich nicht statistisch signifikant sein wird. In der Tat ist der *t*-Wert von $-1,035$ fern der statistischen Signifikanz. (Abweichungen der statistischen Werte im Vergleich zur Beispielrechnung im Buch sind auf Rundungsfehler zurückzuführen. Das Vorzeichen des *t*-Werts hängt davon ab, ob man den zweiten Messzeitpunkt vom ersten subtrahiert oder umgekehrt. Für die Bewertung der Signifikanz hat das Vorzeichen keine Bedeutung, da die *t*-Verteilung symmetrisch ist, siehe Kapitel 3.1)

Beachten Sie schließlich bei den Freiheitsgraden, dass hier – anders als beim *t*-Test für unabhängige Stichproben – $N-1$ Freiheitsgrade zur Verfügung stehen. Da jede Person Werte für beide Testvariablen liefert, können $N-1$ Differenzen frei variieren. Nur die letzte ist festgelegt. Beim *t*-Test für unabhängige Stichproben gilt dieselbe Logik, aber innerhalb jeder der beiden miteinander

<https://lehrbuch-psychologie.springer.com/content/zusatztexte-mit-anleitungen-zu-spss-r-und-gpower-sowie-datensätze>

Aus: Rasch, Friese, Hofmann & Naumann (2021). *Quantitative Methoden. Band 1*, 5. Auflage. Heidelberg: Springer.

© Springer-Verlag GmbH Deutschland, ein Teil von Springer Nature

R-Ergänzungen

Rasch, Frieze, Hofmann & Naumann (2021). *Quantitative Methoden. Band 1* (5. Auflage). Heidelberg: Springer.

vergleichenen Gruppen. Dies führt beim t -Test für unabhängige Stichproben zu $N-2$ Freiheitsgraden (siehe oben sowie Kapitel 3).

<https://lehrbuch-psychologie.springer.com/content/zusatztexte-mit-anleitungen-zu-spss-r-und-gpower-sowie-datensätze>

Aus: Rasch, Frieze, Hofmann & Naumann (2021). *Quantitative Methoden. Band 1*, 5. Auflage. Heidelberg: Springer.

© Springer-Verlag GmbH Deutschland, ein Teil von Springer Nature

Durchführung eines t -Tests für eine Stichprobe

Mit dem t -Test für eine Stichprobe lässt sich der Mittelwert einer Variable gegen einen bestimmten Testwert vergleichen. Eine Frage für dieses Verfahren in Verbindung mit dem Gedächtnisexperiment könnte z.B. lauten: Haben die Versuchspersonen im Schnitt signifikant mehr als 9 Adjektive erinnert?

Um einen t -Test für eine Stichprobe durchzuführen lesen Sie zunächst den normalen Beispieldatensatz ein.

```
library(foreign)
beispiel <- read.spss("Beispieldatensatz.sav", to.data.frame = TRUE)
```

Für den t -Test benötigen Sie wieder die Funktion `t.test()`. Geben Sie als erstes die interessierende Variable (in diesem Fall Gesamtzahl erinnerter Adjektive“) an und spezifizieren Sie mit dem Argument `mu` den Wert, gegen den der Mittelwert der Variable getestet werden soll

```
t.ein <- t.test(beispiel$ges, mu = 9)
```

Geben Sie `t.ein`, um folgenden Output zu erhalten:

```
One Sample t-test

data:  beispiel$ges
t = 3.0099, df = 149, p-value = 0.00307
alternative hypothesis: true mean is not equal to 9
95 percent confidence interval:
 9.368676 10.777991
sample estimates:
mean of x
10.07333
```

Im ersten Teil des Outputs erkennen Sie den Mittelwert der Variable sowie dessen Standardabweichung und den Standardfehler des Mittelwertes. Mittelwert und Standardabweichung sollten Ihnen aus Kapitel 1 vertraut vorkommen.

In dem Output erkennen Sie den Mittelwert der Variable und gegen welchen Testwert der Mittelwert verglichen wurde und präsentiert den resultierenden t -Wert (für dessen Berechnung siehe Kapitel 3.5), die Freiheitsgrade sowie das Konfidenzintervall der Differenz. Dieses liegt ausschließlich im positiven Bereich und erstreckt sich von 9,37 bis 10,78. Der t -Wert von 3,01 ist hochsignifikant. Die Versuchspersonen haben also signifikant mehr als neun Adjektive erinnert.

<https://lehrbuch-psychologie.springer.com/content/zusatztexte-mit-anleitungen-zu-spss-r-und-gpower-sowie-datensätze>

Aus: Rasch, Friese, Hofmann & Naumann (2021). *Quantitative Methoden. Band 1*, 5. Auflage. Heidelberg: Springer.

© Springer-Verlag GmbH Deutschland, ein Teil von Springer Nature

Vertiefung:

Vergleich *t*-Test für unabhängige und abhängige Stichproben in R

Anmerkungen: Bitte lesen Sie den Datensatz „Messwiederholung.sav“ ein.

```
library(foreign)
messw <- read.spss("Messwiederholung.sav", to.data.frame = TRUE)
```

Für die Bearbeitung des folgenden Abschnitts ist ein grundlegendes Verständnis des in Kapitel 4 behandelten Konzeptes Zusammenhang/Korrelation hilfreich.

Ein *t*-Test für abhängige Stichproben ist in der Lage, Daten zu analysieren, die nicht unabhängig sind. Das heißt, dass die Daten in irgendeiner Weise miteinander zusammenhängen. Z.B. könnte es sein, dass eine Person, die zu Zeitpunkt 1 bei einer Aufgabe besonders gut abschneidet, auch zu Zeitpunkt 2 erfolgreicher ist als viele andere Versuchsteilnehmer. In diesem Fall sagt also der Testwert von Zeitpunkt 1 auch etwas über den Erfolg an Zeitpunkt 2 aus. Mit anderen Worten: Die Testwertreihen sind korreliert (Kapitel 4).

Mit derselben Logik können Sie die Einstellungsmessung gegenüber zwei politischen Parteien betrachten, sagen wir der SPD und den Grünen. Wenn eine Person Fragen zu beiden Parteien beantwortet, kann es gut sein, dass die Antworten zu beiden Fragen von derselben Eigenschaft „politische Einstellung“ einer Person geprägt werden. Bei unabhängigen Stichproben ist dies nicht möglich, da jede Person nur *entweder* ihre Einstellung zur SPD *oder* zu den Grünen mitteilen würde (siehe Datei „Kapitel 3 – R_GPower_Ergaenzungen.pdf“).

Der zentrale Unterschied zwischen den beiden *t*-Tests ist also die Korreliertheit der Datenreihen im Fall abhängiger Stichproben im Gegensatz zur Unkorreliertheit bei unabhängigen Stichproben. Diese Aussage impliziert, dass die beiden Verfahren zu identischen Ergebnissen kommen sollten, wenn die Datenreihen trotz theoretischer Abhängigkeit in der Praxis unkorreliert sein sollten; ein Ergebnis, das theoretisch möglich ist. Diese Annahme werden wir in diesem Abschnitt überprüfen.

Wie unterscheidet sich die Datenstruktur für beide Verfahren in R? Vergewähren Sie sich, wie die Daten in dem Gedächtnisexperiment in R angeordnet sind: Es gibt die Variable *bed*, die anzeigt, ob eine Person strukturell, bildhaft oder emotional verarbeitet hat. Jede Zeile gibt die Daten einer Person wieder, so dass die ersten 50 Zeilen Personen mit struktureller Verarbeitung beschreiben usw. Zusätzlich gibt es die Variable *ges*, welche die Gesamtzahl erinnerter Adjektive für jede Person anzeigt. In einem *t*-Test für unabhängige Stichproben ist dies die abhängige Variable, während die Bedingung die Gruppierungsvariable darstellt (siehe Graphik).

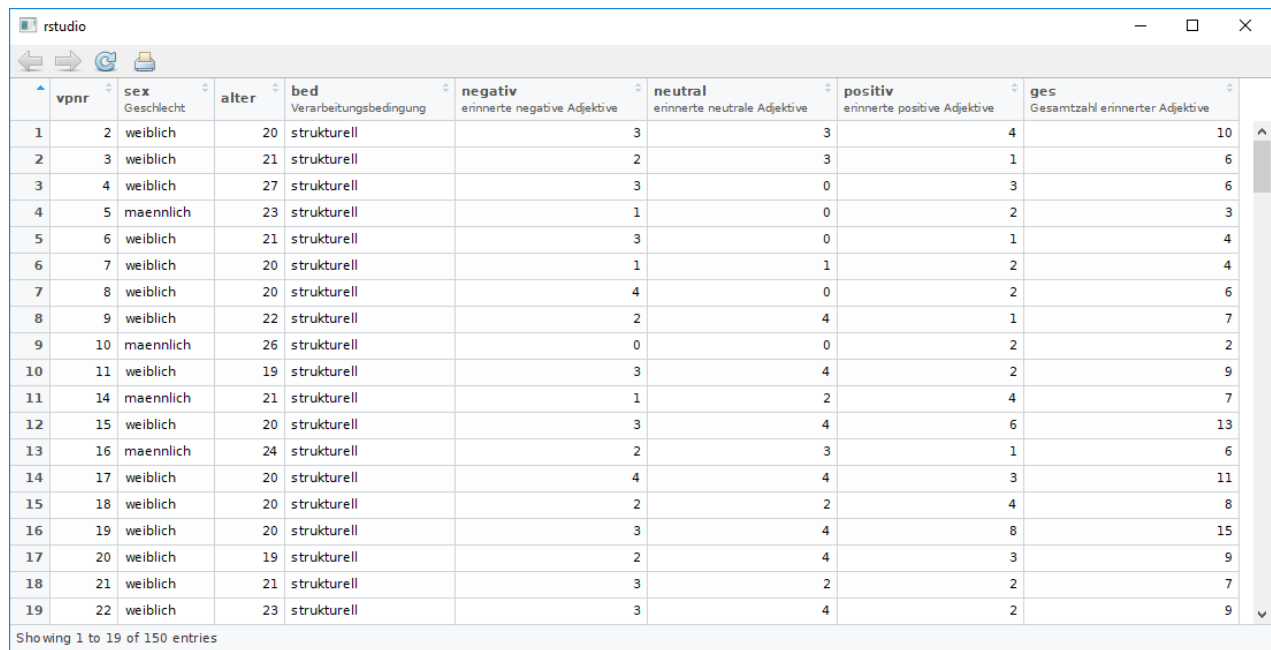
<https://lehrbuch-psychologie.springer.com/content/zusatztexte-mit-anleitungen-zu-spss-r-und-gpower-sowie-datensätze>

Aus: Rasch, Friese, Hofmann & Naumann (2021). *Quantitative Methoden. Band 1*, 5. Auflage. Heidelberg: Springer.

© Springer-Verlag GmbH Deutschland, ein Teil von Springer Nature

R-Ergänzungen

Rasch, Friese, Hofmann & Naumann (2021). *Quantitative Methoden. Band 1* (5. Auflage). Heidelberg: Springer.



vpnr	sex Geschlecht	alter	bed Verarbeitungsbedingung	negativ erinnerte negative Adjektive	neutral erinnerte neutrale Adjektive	positiv erinnerte positive Adjektive	ges Gesamtzahl erinnertes Adjektive	
1	2 weiblich	20	strukturell		3	3	4	10
2	3 weiblich	21	strukturell		2	3	1	6
3	4 weiblich	27	strukturell		3	0	3	6
4	5 maennlich	23	strukturell		1	0	2	3
5	6 weiblich	21	strukturell		3	0	1	4
6	7 weiblich	20	strukturell		1	1	2	4
7	8 weiblich	20	strukturell		4	0	2	6
8	9 weiblich	22	strukturell		2	4	1	7
9	10 maennlich	26	strukturell		0	0	2	2
10	11 weiblich	19	strukturell		3	4	2	9
11	14 maennlich	21	strukturell		1	2	4	7
12	15 weiblich	20	strukturell		3	4	6	13
13	16 maennlich	24	strukturell		2	3	1	6
14	17 weiblich	20	strukturell		4	4	3	11
15	18 weiblich	20	strukturell		2	2	4	8
16	19 weiblich	20	strukturell		3	4	8	15
17	20 weiblich	19	strukturell		2	4	3	9
18	21 weiblich	21	strukturell		3	2	2	7
19	22 weiblich	23	strukturell		3	4	2	9

Showing 1 to 19 of 150 entries

Anders ist die Anordnung bei abhängigen Stichproben. In diesem Fall gibt es keine Gruppierungsvariable wie die Bedingung im Fall von unabhängigen Stichproben. Oben haben wir die mehrfache Messung einer motorischen Aufgabe betrachtet. Alle Versuchspersonen haben zwei Mal an der Aufgabe teilgenommen. Deshalb gibt es dafür zwei eigene Spalten, wie in der folgenden Graphik noch einmal verdeutlicht.

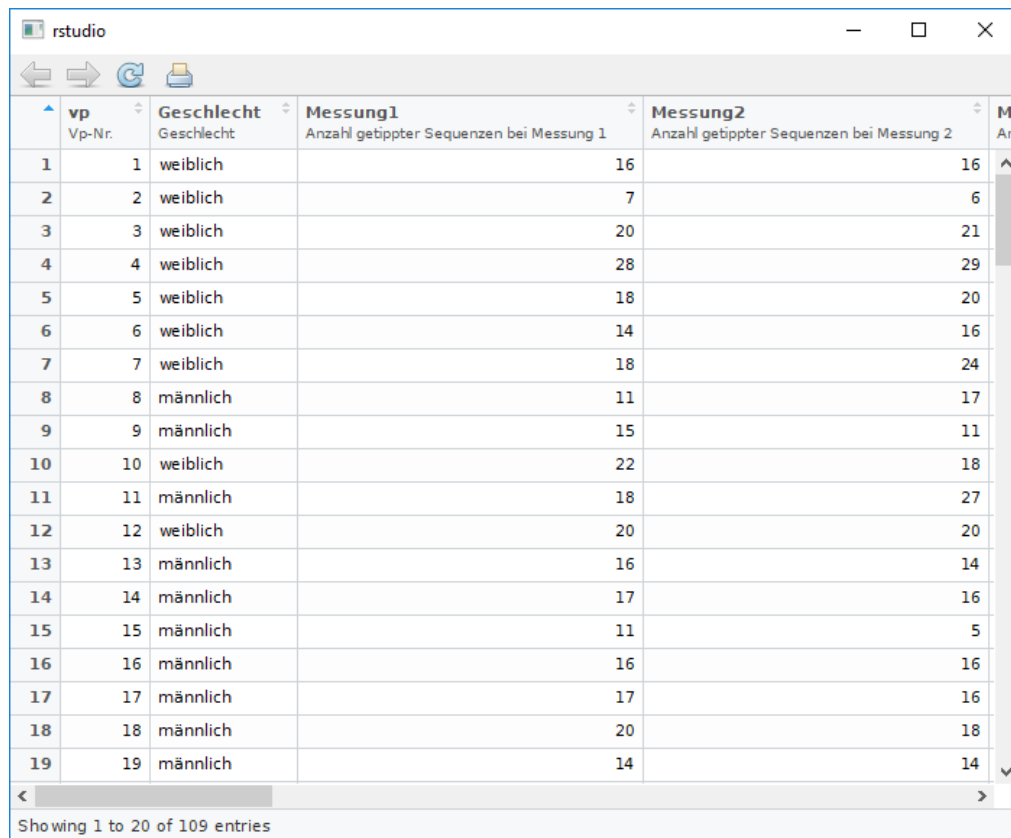
<https://lehrbuch-psychologie.springer.com/content/zusatztexte-mit-anleitungen-zu-spss-r-und-gpower-sowie-datensätze>

Aus: Rasch, Friese, Hofmann & Naumann (2021). *Quantitative Methoden. Band 1*, 5. Auflage. Heidelberg: Springer.

© Springer-Verlag GmbH Deutschland, ein Teil von Springer Nature

R-Ergänzungen

Rasch, Frieze, Hofmann & Naumann (2021). *Quantitative Methoden. Band 1* (5. Auflage). Heidelberg: Springer.



vp Vp-Nr.	Geschlecht Geschlecht	Messung1 Anzahl getippter Sequenzen bei Messung 1	Messung2 Anzahl getippter Sequenzen bei Messung 2
1	weiblich	16	16
2	weiblich	7	6
3	weiblich	20	21
4	weiblich	28	29
5	weiblich	18	20
6	weiblich	14	16
7	weiblich	18	24
8	männlich	11	17
9	männlich	15	11
10	weiblich	22	18
11	männlich	18	27
12	weiblich	20	20
13	männlich	16	14
14	männlich	17	16
15	männlich	11	5
16	männlich	16	16
17	männlich	17	16
18	männlich	20	18
19	männlich	14	14

Zurück zum Vergleich von unabhängigen und abhängigen Stichproben, wenn die abhängigen Stichproben unkorreliert sind. Im „Messwiederholung.sav“ finden Sie die Variablen Frage1 und Frage2. Stellen Sie sich vor, diese beiden Fragen seien Einstellungsangaben zu zwei unterschiedlichen Themen. Die beiden Variablen sind künstlich so gestaltet, dass sie unkorreliert sind, d.h. die Ausprägung auf der Variable Frage1 sagt nichts über die Ausprägung auf der Variable Frage2 einer Person aus. Sie können diese Annahme mit der Funktion `cor()` überprüfen. Geben Sie die beiden Variablen ein. Das Argument `use` spezifiziert, wie mit fehlenden Werten in den Daten umgegangen werden soll. Das Attribut `pairwise.complete.obs` gibt an, dass Fälle mit fehlenden Werten bei den beiden Variablen ausgeschlossen werden sollen.

```
cor(messw$Frage1, messw$Frage2, use = "pairwise.complete.obs")  
[1] 0.000000000000001140113
```

Die minimale Abweichung von Null können wir für unsere Zwecke vernachlässigen. Nehmen wir nun an, die beiden Einstellungsfragen seien nicht denselben Personen gestellt worden, sondern unterschiedlichen. Nehmen wir weiterhin an, dass diese zweite Gruppe von Personen die Werte abgegeben hat, die in der Variable Frage2 abgetragen sind. Die Variable „av“ gibt dieses Gedankenexperiment wieder. Die ersten 36 Werte sind identisch mit der Variable Frage1, während die zweiten 36 Fälle identisch sind mit den Werten der Variable Frage2 (die Werte der Variable Frage2 wurden also einfach dorthin kopiert). Zusätzlich dient die Variable „bed“ als Gruppenvariable, um die Personen zu unterscheiden, welche die erste Frage gestellt bekommen

<https://lehrbuch-psychologie.springer.com/content/zusatztexte-mit-anleitungen-zu-spss-r-und-gpower-sowie-datensätze>

Aus: Rasch, Frieze, Hofmann & Naumann (2021). *Quantitative Methoden. Band 1*, 5. Auflage. Heidelberg: Springer.

© Springer-Verlag GmbH Deutschland, ein Teil von Springer Nature

haben und welche die zweite (analog zur Variable Verarbeitungsbedingung im Gedächtnisexperiment).

Wir haben also folgende Situation: Identische Daten sind einmal so angeordnet, dass wir sie mit einem t -Test für unabhängige Stichproben auswerten können und einmal so, dass wir sie mit einem t -Test für abhängige Stichproben auswerten können. Da die Messwertreihen unkorreliert sind, entfällt der Unterschied zwischen den beiden Arten des t -Tests. Wir sollten folglich in beiden Analysen identische Ergebnisse erzielen.

Bevor wir die Analysen starten, müssen wir noch eine vorbereitende Maßnahme treffen: Für spätere Abschnitte sind an dieser Stelle im Datensatz noch weitere Daten eingerichtet. Wir beschränken uns aber momentan auf die Daten der Gruppen 1 und 2. Deshalb müssen wir einen Teildatensatz erstellen, der alle weiteren, momentan nicht benötigten Fälle ausschließt. Hierfür brauchen wir die Funktion `subset()` und wählen die Fälle aus, die entweder in der Bedingung 1 oder 2 sind:

```
gr1.2 <- subset(messw, messw$bed == 1 | messw$bed == 2)
```

Der senkrechte Strich „|“ bedeutet „oder“. Wenn Sie nun in die Konsole `view(gr1.2)` eingeben, dann sehen Sie, dass der Datensatz nur noch die Fälle enthält, die entweder in der Bedingung 1 oder in der Bedingung 2 sind.

Beginnen wir mit dem t -Test für unabhängige Stichproben. Sie wählen „av“ als abhängige Variable und „bed“ als Gruppierungsvariable des Teildatensatzes.

```
gr1.2 <- subset(messw, messw$bed == 1 | messw$bed == 2)
t.unab.gr1.2 <- t.test(gr1.2$av ~ gr1.2$bed, var.equal = TRUE)
```

Sie erhalten folgenden Output, wenn Sie `t.unab.gr1.2` in die Konsole eingeben:

```
Two Sample t-test

data:  gr1.2$av by gr1.2$bed
t = -0.68847, df = 70, p-value = 0.4934
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 -2.815302  1.370413
sample estimates:
mean in group 1 mean in group 2
    16.55556      17.27800
```

Die beiden Mittelwerte unterscheiden sich um 0,72244 Einheiten auf der abhängigen Variablen. Der Unterschied ist bei weitem nicht signifikant. Wie sieht das Ergebnis aus, wenn wir dieselben Daten mit einem t -Test für abhängige Stichproben analysieren?

Wählen Sie `Frage1` und `Frage2` als zu vergleichende Variablen aus. Da aus der Bedingung 2 kein Fall Werte auf einer dieser beiden Variablen aufweist, brauchen wir den Filter nicht zu verändern. Es werden nur die 36 Fälle eingeschlossen, die Werte auf den interessierenden Variablen aufweisen.

<https://lehrbuch-psychologie.springer.com/content/zusatztexte-mit-anleitungen-zu-spss-r-und-gpower-sowie-datensätze>

Aus: Rasch, Friese, Hofmann & Naumann (2021). *Quantitative Methoden. Band 1*, 5. Auflage. Heidelberg: Springer.

© Springer-Verlag GmbH Deutschland, ein Teil von Springer Nature

```
t.abh.gr1.2 <- t.test(gr1.2$Frage1, gr1.2$Frage2, paired = TRUE)
```

Sie erhalten folgenden Output:

```
Paired t-test

data:  gr1.2$Frage1 and gr1.2$Frage2
t = -0.68847, df = 35, p-value = 0.4957
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 -2.852733  1.407845
sample estimates:
mean of the differences
      -0.7224444
```

Sie sehen, dass die Mittelwerte der beiden Fragen den Mittelwerten der Gruppen 1 und 2 beim t -Test für unabhängige Stichproben entsprechen. Sie führen zur identischen Differenz von 0,72244. Auch der t -Wert von 0,688 ist identisch, ebenso wie der Signifikanzwert (abgesehen von geringfügigen Abweichungen, die auf Grund der unterschiedlichen Freiheitsgradanzahl zustande kommen, vgl. Kap. 3.1.4). Die theoretische Annahme hat sich also bestätigt: Bei unkorrelierten Messwertreihen führen der t -Test für unabhängige und abhängige Stichproben zu identischen Ergebnissen.

<https://lehrbuch-psychologie.springer.com/content/zusatztexte-mit-anleitungen-zu-spss-r-und-gpower-sowie-datensätze>

Aus: Rasch, Frieze, Hofmann & Naumann (2021). *Quantitative Methoden. Band 1*, 5. Auflage. Heidelberg: Springer.

© Springer-Verlag GmbH Deutschland, ein Teil von Springer Nature

Literatur

Delacre, M., Lakens, D., & Leys, C. (2017). Why psychologists should by default use Welch's t-test instead of Student's t-test. *International Review of Social Psychology*, 30, 92-101. <https://doi.org/10.5334/irsp.82>

<https://lehrbuch-psychologie.springer.com/content/zusatztexte-mit-anleitungen-zu-spss-r-und-gpower-sowie-datensätze>

Aus: Rasch, Friese, Hofmann & Naumann (2021). *Quantitative Methoden. Band 1*, 5. Auflage. Heidelberg: Springer.

© Springer-Verlag GmbH Deutschland, ein Teil von Springer Nature